IN THE CLAIMS

1. (Previously Presented) A computer-implemented method in which a computer system initiates execution of software instructions stored in memory, the computer-implemented method comprising:

accepting text spellings of training words in a plurality of sets of training words, each set corresponding to a different one of a plurality of languages;

for each of the sets of training words in the plurality, receiving pronunciations for the training words in the set, the pronunciations being characteristic of native speakers of the language of the set, the pronunciations also being in terms of subword units at least some of which are common to two or more of the languages; and

training a single pronunciation estimator using data comprising the text spellings and the pronunciations of the training words; and

calculating a single acoustic subword model for each subword unit, based on the pronunciations in the plurality of sets of training words, by mixing distributions of acoustic parameters representing the sounds of the subword unit in multiple languages when a subword unit is common to two or more languages.

2. (Previously Presented) The computer-implemented method of claim 1 further comprising:

accepting a plurality of sets of utterances, each set corresponding to a different one of the plurality of languages, the utterances in each set being spoken by the native speakers of the language of each set; and

training a set of acoustic models for the subword units using the accepted sets of utterances and pronunciations estimated by the single pronunciation estimator from text representations of the training utterances.

3. (Previously Presented) The computer-implemented method of claim 1, wherein a first training word in a first set in the plurality corresponds to a first language and a second training word in a second set corresponds to a second language, the first and second training words having identical text spellings, the received pronunciations for the first and second training words being different.

4. (Previously Presented) The computer-implemented method of claim 3, wherein utterances of the first and the second training words are used to train a common subset of subword units.

5. (Previously Presented) The computer-implemented method of claim 1, wherein the single pronunciation estimator uses a decision tree to map letters of the text spellings to pronunciation subword units.

6. (Previously Presented) The computer-implemented method of claim 1, where training the single pronunciation estimator further comprises: forming, from sequences of letters of each training word's textual spelling and the corresponding grouping of subword units of the pronunciation, a letter to subword mapping for each training word; and training the single pronunciation estimator using the letter-to-subword mappings.

7. (Previously Presented) The computer-implemented method of claim 6, wherein training the single pronunciation estimator and training the acoustic models is executed by a nonportable programmable device.

8. (Previously Presented) The computer-implemented method of claim 1 further comprising:

generating, for each word in a list of words to be recognized, an acoustic word model, the generating comprising generating a grouping of subword units representing a pronunciation of the word to be recognized using the single pronunciation estimator.

9. (Previously Presented) The computer-implemented method of claim 8, wherein the grouping of subword units is a linear sequence of subword units.

10. (Previously Presented) The computer-implemented method of claim 9, wherein the grouping of the acoustic subword models is a linear sequence of acoustic subword models.

11. (Previously Presented) The computer-implemented method of claim 8, wherein the subword units are phonemes.

12. (Previously Presented) The computer-implemented method of claim 8, wherein the grouping of subwords is a network, and the network represents two pronunciations of a

word, the two pronunciations being representative of utterances of native speakers of two languages.

13. (Previously Presented) The computer-implemented method of claim 8 further comprising:

processing an utterance; and

scoring matches between the processed utterance and the acoustic word models.

14. (Previously Presented) The computer-implemented method of claim 13, wherein generating the acoustic word model, processing the utterance, and scoring matches is executed by a portable programmable device.

15. (Previously Presented) The computer-implemented method of claim 14, wherein the portable programmable device is a cellphone.

16. (Previously Presented) The computer-implemented method of claim 13, wherein the utterance is spoken by a native speaker of one of the plurality of languages.

17. (Previously Presented) The computer-implemented method of claim 14, wherein the utterance is spoken by a native speaker of a language other than the plurality of languages, the language having similar sounds and similar letter to sounds rules as a language from the plurality of languages.

18. (Previously Presented) A computer-implemented method in which a computer system initiates execution of software instructions stored in memory for recognizing words spoken by native speakers of multiple languages, the computer-implemented method comprising:

generating a set of estimated pronunciations, using a single pronunciation estimator, from text spellings of a set of acoustic training words, each pronunciation comprising a grouping of subword units, the set of acoustic training words comprising at least a first word and a second word, the first and second words having identical text spelling, the first word having a pronunciation based on utterances of native speakers of a first language, the second word having a pronunciation based on utterances of native speakers of a second language;

mapping sequences of sound associated with utterances of each of the acoustic training words against the estimated pronunciation associated with each of the acoustic training words; and

using the mapping of sequences of sound to estimated pronunciations to generate a single acoustic subword model for each of the subword units in the grouping of subwords, by mixing distributions of acoustic parameters representing the sounds of the subword unit in multiple languages when a subword unit is common to two or more languages, the acoustic subword model comprising a sound model and a subword unit.

19. (Previously Presented) A computer-implemented method in which a computer system initiates execution of software instructions stored in memory for multilingual speech recognition, the computer-implemented method comprising:

accepting a recognition vocabulary that includes words from multiple languages;

determining a pronunciation of each of the words in the recognition vocabulary using a pronunciation estimator that is common to the multiple languages;

determining an acoustic word model for each of the words in the recognition vocabulary by mapping subword units in the estimated pronunciation to acoustic subword models, at least some of which comprise a mix of distributions of acoustic parameters representing the sounds of the subword unit in multiple languages, and combining the acoustic subword models; and

configuring a speech recognizer using the determined acoustic word models of the words in the recognition vocabulary.

20. (Previously Presented) The computer-implemented method of claim 19 further comprising:

accepting a training vocabulary that comprises words from multiple languages;

determining a pronunciation of each of the words in the training vocabulary using the pronunciation estimator that is common to the multiple languages;

configuring the speech recognizer using parameters estimated using the determined pronunciations of the words in the training vocabulary; and

recognizing utterances using the configured speech recognizer.

21. (Previously Presented) A computer program product, tangibly embodied in a storage medium, the computer program product being operable to cause data processing apparatus to:

accept text spellings of training words in a plurality of sets of training words, each set corresponding to a different one of a plurality of languages;

for each of the sets of training words in the plurality, receive pronunciations for the training words in the set, the pronunciations being characteristic of native speakers of the language of the set, the pronunciations also being in terms of subword units at least some of which are common to two or more of the languages;

train a pronunciation estimator using data comprising the text spellings and the pronunciations of the training words; and

calculating a single acoustic subword model for each subword unit, based on the pronunciations in the plurality of sets of training words, by mixing distributions of acoustic parameters representing the sounds of the subword unit in multiple languages when a subword unit is common to two or more languages.

22. (Original) The computer program product of claim 21, the computer program product being further operable to cause the data processing apparatus to:

accept a plurality of sets of utterances, each set corresponding to a different one of the plurality of languages, the utterances in each set being spoken by the native speakers of the language of each set; and

train a set of acoustic models for the subword units using the accepted sets of utterances and pronunciations estimated by the single pronunciation estimator from text

representations of the training utterances.

23. (Original) The computer program product of claim 22, wherein a first training word in a first set in the plurality corresponds to a first language and a second training word in a second set corresponds to a second language, the first and second training words having identical text spellings, the received pronunciations for the first and second training words being different.

24. (Original) The computer program product of claim 23, wherein utterances of the first and the second training words are used to train a common subset of subword units.

25. (Original) The computer program product of claim 21, wherein the single pronunciation estimator uses a decision tree to map letters of the text spellings to pronunciation subword units.

26. (Original) The computer program product of claim 21, wherein training the single pronunciation estimator further comprises:

form, from sequences of letters of each training word's textual spelling and the corresponding grouping of subword units of the pronunciation, a letter to subword mapping for each training word; and

train the single pronunciation estimator using the letter-to-subword mappings.

27. (Original) The computer program product of claim 22, wherein training the single

pronunciation estimator and training the acoustic models is executed by a nonportable programmable device.

28. (Original) The computer program product of claim 22, the computer program product being further operable to cause the data processing apparatus to:

generate, for each word in a list of words to be recognized, an acoustic word model, the generating comprising generating a grouping of subword units representing a pronunciation of the word to be recognized using the single pronunciation estimator.

29. (Original) The computer program product of claim 28 wherein the grouping of subword units is a linear sequence of subword units.

30. (Original) The computer program product of claim 29, wherein the grouping of the acoustic subword models is a linear sequence of acoustic subword models.

31. (Original) The computer program product of claim 28, wherein the subword units are phonemes.

32. (Original) The computer program product of claim 28, wherein the grouping of subwords is a network, and the network represents two pronunciations of a word, the two pronunciations being representative of utterances of native speakers of two languages.

33. (Original) The computer program product of claim 28, the computer program product being further operable to cause the data processing apparatus to:

process an utterance; and

score matches between the processed utterance and the acoustic word models.

34. (Original) The computer program product of claim 33, wherein generating the acoustic word model, processing the utterance, and scoring matches is executed by a portable programmable device.

35. (Original) The computer program product of claim 34, wherein the portable programmable device is a cellphone.

36. (Original) The computer program product of claim 33, wherein the utterance is spoken by a native speaker of one of the plurality of languages.

37. (Original) The computer program product of claim 35, wherein the utterance is spoken by a native speaker of a language other than the plurality of languages, the language having similar sounds and similar letter to sounds rules as a language from the plurality of languages.

38. (Currently Amended) A computer program product, tangibly embodied in a storage medium, for recognizing words spoken by native speakers of multiple languages, the computer program product being operable to cause data processing apparatus to:

generate a set of estimated pronunciations, using a single pronunciation estimator, from text spellings of a set of acoustic training words, each pronunciation comprising a grouping of subword units, the set of acoustic training words comprising at least a first word and a second word, the first and second words having identical text spelling, the first word having a pronunciation based on utterances of native speakers of a first language, the second word having a pronunciation based on utterances of native speakers of a second language;

map sequences of sound associated with utterances of each of the acoustic training words against the estimated pronunciation associated with each of the acoustic training words; and

use the mapping of sequences of sound to estimated pronunciations to generate a single acoustic subword model for each of the subword units in the grouping of subwords, by mixing distributions of acoustic parameters representing the sounds of the subword unit in multiple languages when a subword model comprising a sound model and a subword unit.

39. (Currently Amended) A computer program product, tangibly embodied in a storage medium, for multilingual speech recognition, the computer program product being operable to cause data processing apparatus to:

accept a recognition vocabulary that includes words from multiple languages;

determine a pronunciation of each of the words in the recognition vocabulary using a pronunciation estimator that is common to the multiple languages;

determining an acoustic word model for each of the words in the recognition vocabulary by mapping subword units in the estimated pronunciation to acoustic subword models, at least some of which comprise a mix of distributions of acoustic parameters representing the sounds of the subword unit in multiple languages, and combining the acoustic subword models; and

configure a speech recognizer using the determined acoustic word models of the words in the recognition vocabulary.

40. (Previously Presented) The computer program product of claim 39, the computer program product being further operable to cause data processing apparatus to:

accept a training vocabulary that comprises words from multiple languages;

determine a pronunciation of each of the words in the training vocabulary using the pronunciation estimator that is common to the multiple languages;

configure the speech recognizer using parameters estimated using the determined pronunciations of the words in the training vocabulary; and

recognize utterances using the configured speech recognizer.

41. (Currently Amended) ~~An apparatus~~A computer system comprising:

a processor;

a memory coupled to the processor, the memory storing instructions that when executed by the processor cause the system to perform the operations of:

~~means for~~ accepting text spellings of training words in a plurality of sets of training words, each set corresponding to a different one of a plurality of languages;

~~means for~~ receiving, for each of the sets of training words in the plurality, pronunciations for the training words in the set, the pronunciations being characteristic of native speakers of the language of the set, the pronunciations also being in terms of subword units at least some of which are common to two or more of the languages;

~~means for~~ training a single pronunciation estimator using data comprising the text spellings and the pronunciations of the training words; and

~~means for~~ calculating a single acoustic subword model for each subword unit, based on pronunciations in the plurality of sets of training words, by fixing distributions of acoustic parameters representing the sounds of the subword unit in multiple languages when a subword unit is common to two or more languages.

42. (Currently Amended) The ~~apparatus~~ computer system of claim 41 ~~further comprising~~ the memory storing further instructions that when executed by the processor causes the system to perform the operations of:

~~means for~~ accepting a plurality of sets of utterances, each set corresponding to a different one of the plurality of languages, the utterances in each set being spoken by the native speakers of the language of each set; and

~~means for~~ training a set of acoustic models for the subword units using the accepted sets of utterances and pronunciations estimated by the single pronunciation

estimator from text representations of the training utterances.

43. (Currently Amended) The ~~apparatus~~ computer system of claim 42 ~~further comprising~~ the memory storing further instructions that when executed by the processor causes the system to perform the operations of:

~~a means for~~ generating, for each word in a list of words to be recognized, an acoustic word model, the generating comprising generating a grouping of subword units representing a pronunciation of the word to be recognized using the single pronunciation estimator.

44. (Currently Amended) The ~~apparatus~~ computer system of claim 43 ~~further comprising~~ the memory storing further instructions that when executed by the processor causes the system to perform the operations of:

~~means for~~ processing an utterance; and

~~means for~~ scoring matches between the processed utterance and the acoustic word models.

45. (Currently Amended) ~~An apparatus~~A computer system for recognizing words spoken by native speakers of multiple languages, the ~~apparatus~~ computer system comprising:

a processor;

a memory coupled to the processor, the memory storing instructions that when executed by the processor cause the system to perform the operations of:

~~a means for~~ generating a set of estimated pronunciations, using a pronunciation estimator, from text spellings of a set of acoustic training words, each pronunciation comprising a grouping of subword units, the set of acoustic training words comprising at least a first word and a second word, the first and second words having identical text spelling, the first word having a pronunciation based on utterances of native speakers of a first language, the second word having a pronunciation based on utterances of native speakers of a second language;

~~means for~~ mapping sequences of sound associated with utterances of each of the acoustic training words against the estimated pronunciation associated with each of the acoustic training words; and

~~means for~~ using the mapping of sequences of sound to estimated pronunciations to generate a single acoustic subword model for each of the subword units in the grouping of subwords, by mixing distributions of acoustic parameters representing the sounds of the subword unit in multiple languages when a subword unit is common to two or more languages, the acoustic subword model comprising a sound model and a subword unit.

46. (Currently Amended) ~~An apparatus~~A computer system for multilingual speech recognition, the ~~apparatus~~ computer system comprising:

a processor;

a memory coupled to the processor, the memory storing instructions that when executed by the processor cause the system to perform the operations of:

~~means for~~ accepting a recognition vocabulary that includes words from multiple languages;

~~means for~~ determining a pronunciation of each of the words in the recognition vocabulary using a pronunciation estimator that is common to the multiple languages;

~~means for~~ determining a pronunciation of each of the words in the recognition vocabulary using a pronunciation estimator that is common to the multiple languages;

~~means for~~ determining an acoustic word model for each of the words in the recognition vocabulary by mapping subword units in the estimated pronunciation to acoustic subword models, at least some of which comprise a mix of distributions of acoustic parameters representing the sounds of the subword unit in multiple languages, and combining the acoustic subword models; and

~~means for~~ configuring a speech recognizer using the determined acoustic words models of the words in the recognition vocabulary.

47. (Previously Presented) The computer-implemented method of claim 1, wherein mixing distributions of acoustic parameters from multiple languages comprises mixing Gaussian probability distributions of acoustic parameters from multiple languages.

48. (Previously Presented) A computer-implemented method in which a computer system initiates execution of software instructions stored in memory, the computer-implemented method comprising:

accepting text spellings of training words in a plurality of sets of training words, each set corresponding to a different one of a plurality of languages;

for each of the sets of training words in the plurality, receiving pronunciations for the training words in the set, the pronunciations being characteristic of native speakers of the language of the set, the pronunciations also being in terms of subword units at least some of which are common to two or more of the languages;

training a pronunciation estimator using data comprising the text spellings and the pronunciations of the training words; and

calculating an acoustic subword model for each subword unit, based on the pronunciations in the plurality of sets of training words, by mixing distributions of acoustic parameters from multiple languages when a subword unit is common to two or more languages, wherein an acoustic subword model for a subword unit that is common to two or more languages comprises a probability distribution that is a weighted blend of probability distributions each corresponding to a different sound associated with the subword unit.